EVE: An Environment for On-board Processing

Steve Tanner, M. Alshayeb, E. Criswell, S. Graves, A. McDowell, and M. McEniry
Information Technology and Systems Center
University of Alabama in Huntsville
S345 Technology Hall
Huntsville, Alabama 35899
stanner@itsc.uah.edu
www.itsc.uah.edu

Abstract - The Information Technology and Systems Center (ITSC) at The University of Alabama in Huntsville (UAH) is investigating and developing an innovative processing system capable of handling the unique constraints and characteristics of the on-board satellite data and information environment. The EnVironmEnt for On-Board Processing (EVE) system will serve as a proof-of-concept of advanced information systems technology for remote sensing platforms. EVE's on-board, real-time processing will provide capabilities focused on the areas of autonomous data mining, classification and feature extraction. These will contribute to Earth Science research applications, including natural hazard detection and prediction, fusion of multi-sensor measurements, intelligent sensor control, and the generation of customized data products for direct distribution to users. EVE is being engineered to provide high performance data processing in a real-time operational environment. A ground-based testbed is being created to provide testing of EVE and associated Earth Science applications in a heterogeneous embedded hardware and software environment. *

INTRODUCTION

In this paper, the Information Technology and Systems Center (ITSC) at The University of Alabama in Huntsville (UAH) presents current efforts to address on-board satellite data processing problems based on the Center's existing custom processing, feature extraction, and data mining technologies. This project, the Environment for On-Board Processing (EVE), benefits from the ITSC's extensive experience with scientific data mining and knowledge discovery. Other current ITSC research projects deal with the distributed and heterogeneous nature of sensor based Earth Science data sets, as well as data integration, data fusion, and high-performance networking. Based on this background and a broad range of research affiliations, the ITSC is investigating, designing and developing a new breed of processing system capable of handling the unique constraints and characteristics of the on-board data and information environment. The final EVE system will be adaptable to new Earth Science measurements, and will enable new information products.

According to NASA's Earth Science Vision, on-board processing will play a significant role in the next generation of Earth Science missions, providing the opportunity for greater flexibility and versatility in measurements of the Earth's systems. Such on-board processing can contribute to many Earth Science research applications, including natural hazard detection and prediction, intelligent sensor control, and the generation of customized data products for direct distribution to users. Ideally, the availability of custom processing, feature extraction, and data mining on-board satellites can allow end users to specify their own data products through the definition of a processing plan. On-board processing will reduce the volume of delivered data since only the data specified by the processing plan is transmitted to the user. This in turn will improve the accessibility and utility of Earth Science data sets, and overcome in part, the autonomous nature of Satellite sensor data.

While the ITSC has made significant progress in custom processing [1,2,3,4] and data mining technology [5,6], the current architectures were not designed to be optimal for on-board processing in an autonomous environment. In particular, the on-board processing environment will include significant hardware constraints, coupled with requirements for processing real-time streams of sensor measurements. Furthermore, the need for fusion of multi-sensor inputs, both on-board a single craft, and ultimately through communication between crafts, is of great importance. Therefore, EVE requires a complete re-engineering of current well-established software, ranging from fundamental changes in basic system architecture through new implementation of processing modules for even greater efficiency. This approach will reduce the risk, cost, and time associated with development of a full suite of on-board processing functionality. Integral to this effort is the simultaneous development of a ground-based testbed, which enables researchers to perform testing and certification in an environment simulating the expected on-board processing environment.

During the latter phases of this research, the ITSC hopes to take advantage of a flight of opportunity, which would allow an actual on-board test of the EVE prototype

* Financial and technical support for this effort is provided by NASA's Earth Science Technology Office.

on an Uninhabited Aerial Vehicle (UAV) experimental flight. These high altitude UAVs fly meteorological instruments for extended time periods (> 18 hrs). By collaborating with NASA scientists who are planning a UAV experiment, the ITSC will define an on-board processing plan that is meaningful in the UAV context. This experiment will complement ground-based testbed research in this proof-of-concept for specialized on-board processing that will include feature extraction and data mining. Goals of such a flight may include some autonomous control and flight planning based upon real-time data mining of sensor input.

The ITSC is embarking on a phased approach for this research, advancing the custom processing and data mining software from NASA's Technology Readiness Level (TRL) 2 to TRL 5 over a three-year period. This means that the system will progress from a purely theoretical research endeavor into a demonstrable system. Much research and development will have to be accomplished to reach these goals.

Description of the EVE Technology

Some of the goals of on-board data processing include developing applications to provide real-time event detection and prediction, feature extraction, intelligent sensor control and reduction in data volume for direct transmission to users. A compact, but flexible and easily configurable processing framework will be required to meet all of these goals. The ITSC's Algorithm Development and Mining (ADaM) system [5], is described in more detail later in this paper. ADaM provides a data processing architecture that supports many applications, including data mining [5,7,8], feature extraction [9,10], classification [11], anomaly detection [12], data subsetting [13], and implementation of multi-discipline scientific algorithms. While the ADaM software cannot be ported directly to an embedded systems environment, the ITSC's extensive experience in the field of data mining, and in the development of system architectures in support of that technology, is being used in the design and development of EVE. The software analysis, engineering, design, and development necessary to introduce this technology into a space platform environment are addressed later in this paper, as well as a testbed for validating the resulting software.

Current Data Mining Technology (ADaM)

The ITSC is an established center for mining and processing Earth Science data, providing data mining services and consultation for a variety of scientific research projects. The foundation for the Center's data mining infrastructure is the ADaM system, originally developed in conjunction with an earlier NRA, "Phenomena Oriented Data Mining" used with Earth Science data sets.

ITSC researchers have continued to enhance and refine the system beyond this initial research to provide an environment that allows varied custom processing functionality, including data mining on ingest, event detection, feature extraction, anomaly detection, subsetting, and classification with improved performance and reliability. Current research includes adapting ADaM for use as a custom data processing service in the Passive Microwave Earth Science Information Partner (PM-ESIP) project. As a ground-based system, ADaM has advanced from Technology Readiness Level 5 to 6 during the course of the PM-ESIP project. This means that the system has progressed from technology development into a tangible and demonstrable system, albeit in a specific but useful domain. This concept of TRL levels is important to EVE because it is a measure of the solid foundation that ADaM provides and from which EVE will be built.

This work has also demonstrated ITSC's ability to customize ADaM components for inclusion into new systems. ADaM is implemented in a Client-Server configuration on Unix and NT workstations and is written entirely in portable C++. Fig. 1 illustrates the ADaM data mining environment, showing the interactions between the ADaM engine, client applications, input data and outputs.

ADaM Architecture. Earth Science data tends to come in many shapes, sizes, formats, scales, and resolutions. This has historically caused time-consuming problems when researchers attempt to integrate new data sets into studies and models. In an attempt to alleviate some of these problems, ADaM was designed to internally handle all data in a common representation regardless of the actual input and resulting output formats. This approach effectively insulates the analysis modules from data set specific concerns, thereby simplifying the development of those modules. Any number of software modules may be chained together in a processing plan with results from one processing operation forming the input to the next operation. Many of ADaM's functional units consist of components that can easily be configured for specific

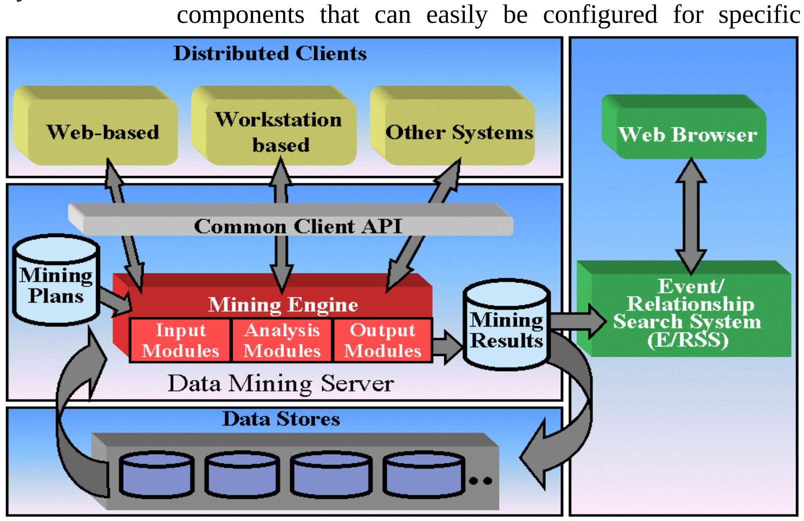

Fig. 1: ADaM Environment for Data Mining and Analysis

applications or replaced with specialized processing if necessary. This component architecture allows ADaM to be highly extensible and adaptive to new configurations, and is also a cornerstone of the EVE project.

ADaM was developed to efficiently handle very large Earth Science data sets in a file-based server environment. Thus, ADaM has evolved characteristics that are incompatible with on-board processing. For example, the software runs on UNIX and Windows operating systems that are not typically employed in on-board environments. Also, ADaM was not designed with real-time processing considerations, nor can it execute multiple processing plans simultaneously. Finally, the complex software system is too resource intensive for the on-board environment.

EnVironmEnt for On-Board Processing (EVE)

ADaM has evolved into a mature system for ground-based data mining functionality, but it was not designed for use in an embedded system environment, as will be required for on-board processing. However, this existing technology, in terms of algorithms and expertise gained, is being applied directly to the development of an on-board processing prototype. But this is not enough, because the EVE on-board data processing configuration is a significantly different architecture. The following sections examine some of the issues that are important for this new design and development effort.

On-Board Processing Constraints. EVE is being designed specifically to support expected on-board environment constraints, such as limited memory and processing capacity, limited power availability, and real-time data streams. These constraints will be simulated in an ITSC testbed so that EVE's architecture can be verified to be compatible with current and projected on-board processing environments. Below are some architectural considerations that are important for EVE design and development:

- Parallel processing capabilities to optimize processing of a large stream of input data.
- Support for multi-processor environments.
- Ability to handle streamed, low level, input data.
- Support for scheduling tasks based on real-time geo-temporal events.
- Support for real-time processing environments.
- Re-configurable processing plans to allow for dynamic system changes.

Fig. 2 illustrates an initial conceptual view of the EVE on-board architecture, in the context of the expected processing environment of a typical satellite. While the original Level 0 data continues to be passed to ground stations intact, the on-board processing system will draw

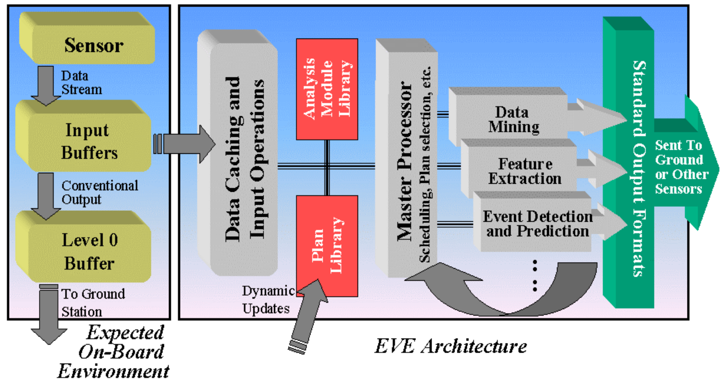

Fig. 2: On-bard EVE Architecture

low level sensor data in real time from the same input buffers. A master processor will assemble analysis modules for specific processing tasks, according to the processing plans. EVE's input operations may include subsetting of the data stream and directed input caching to allow for implementation of scientific algorithms requiring larger data segments for processing. Parallel processing capabilities will allow for applications such as data mining, feature extraction, and event prediction/detection to be applied simultaneously to the stream of sensor readings.

Libraries of analysis modules and processing plan instantiations can be updated to allow for dynamic reconfiguration of the system. For example, detection of specified geo-temporal events can cause an update to the processing schedule. The number of data processing modules supported by this configuration will be determined in the ground-based testbed, based on the capabilities of the planned computer system for a particular space platform. The system will be scalable to provide support ranging from a very minimal platform, in terms of processor speed, memory, disk space, and other factors, to larger and more sophisticated systems when available in the space environment.

Fig 3. illustrates the initial architectural design for the EVE framework. In this framework, users interact with the Processing Plan Editor to chain together their desired set of software modules, much as they do with the current ADaM system. However, several new components and steps come into play.

First, the constraints of the processing environment and the sensor data available must be represented, and may impact the efficacy of a processing plan. This information is stored in an On-board Configuration Library and a Sensor Model Library. Second, the processing plan may now encompass multiple processes running on several on-board and ground-based platforms. Inter-process communications and scheduling tasks must be coordinated as part of the processing plan.

Once a plan has been developed, it can be represented in a standard XML format. This in turn can feed one or more on-board specific cross-compilers that can generate the code required for upload to each affected on-board system.

Software. ITSC is undertaking a phased implementation of the EVE prototype over the course of a 3-year research effort. The first year of research is focused on defining and verifying a solid software architecture for on-board processing. Research and analysis will determine an efficient software component architecture, with a focus on the interfaces between components, and will establish the best methodology for constructing the components on a given hardware architecture. The basic processing framework will be developed and initial processing operations implemented. Like ADaM, EVE consists of software modules that provide specific functionality operating on a common structure. However, EVE's modules are intended to be smaller and more specialized in nature due to the restricted environment involved. The prototype software architecture will be designed specifically so that a substantial library of processing components will be available for easy reconfigurability. This will be accomplished by developing processing components through the use of common building-block interfaces. The team expects current efforts in defining a standard Earth Science Markup Language [14] will make this task significantly easier.

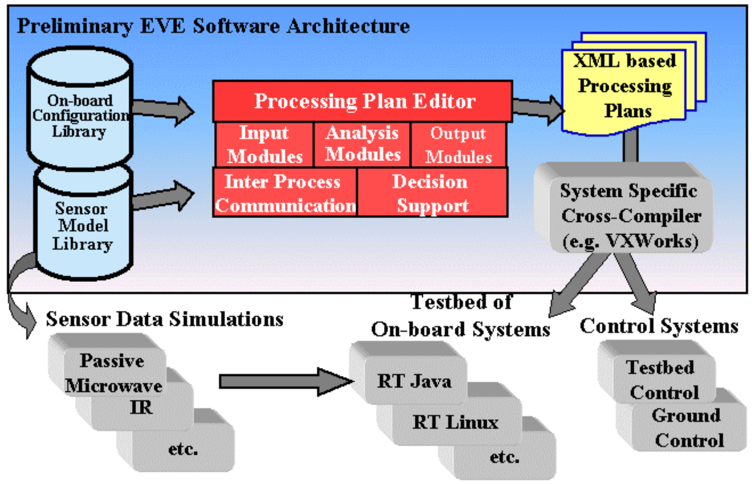

Fig. 3: Preliminary EVE Architecture

Early in the second year, the EVE prototype software will be ported to the real-time embedded system environment of the testbed for validation of the overall software design. Collaborations with Earth Science colleagues will guide the development of additional processing modules to support an on-board processing science scenario. Following the paradigm established by the ADaM mining plans, a processing plan syntax will be developed for configuring EVE's processing modules. This capability will support on-orbit reconfiguration of the processing flow in a mature system. Interfaces between the remote sensing instruments and on-board processing software will also be defined during this period and implemented in the testbed sensor models.

In support of this activity, the team is evaluating both current sensor design labs and embedded systems testbeds including the Space Optics Manufacturing Center and the International Space Station Payload Rack Interface Controller testbed at the NASA Marshall Space Flight Center and the Army's Advanced Simulation Hardware in the Loop testbed at Redstone Arsenal.

During the third research year, data input components will be developed, based on the instrument interfaces defined in Year 2. Software development will also continue with the implementation of additional processing modules and the refinement of the overall system architecture. Mid-way through Year 3, a demonstration of the prototype with the full-scale science processing scenario and appropriate data sets is planned. The final six months of research will focus on thorough testing of the prototype in the testbed with reasonably realistic supporting elements, conforming to target environment and interfaces. Collaboration with NASA Marshall Space Flight Center and others in development of an on-board processing scenario for the Conically-Scanning Two-look Airborne Radiometer (C-STAR) or Advanced Microwave Precipitation Radiometer (AMPR) instruments could lead to a possible flight of opportunity aboard a UAV.

Throughout the research period, ITSC will maintain awareness of emerging data and software standards. ITSC will also explore issues in multiple execution paths and run-time configuration of the software. Although these features are desirable in a mature on-board processing system, they are not required for an initial flight. Therefore, full implementation of such features is outside the scope of this research, but will be considered in the overall architecture design.

On-Board Processing Testbed

EVE is being designed to run in an on-board systems environment. In order to provide adequate ground-based testing, ITSC is constructing an on-board processing testbed. This testbed will allow the immersion of candidate on-board data mining architectures into an environment that closely matches the constraints of actual space-based sensor and processing systems. Initial architectures of the testbed utilize existing resources from the ITSC Data Mining Laboratory. Once fully completed with additional resources, the testbed will be highly adaptable to emerging technologies and will accommodate new hardware, operating systems, special purpose equipment, and unique communications architectures that match satellite systems on the drawing board for the next decade. The ITSC's Internet 2 connection will be utilized for remote access, allowing the potential for instrument teams and the Instrument Incubator Program (IIP) to interface to the testbed. The on-board processing testbed will consist of three major components: the Sensor Simulations; the Embedded On-Board Processing System; and the Testbed Control System.

Sensor Simulations. The close coupling of on-board processing with the acquisition of low-level data will require Sensor Simulations that can deliver accurate and timely data to the EVE prototype. The testbed will have two Sensor Simulation methods. One will use data from existing missions such as the Lightning Image Sensor (LIS), Geosynchronous Operational Environmental Satellites (GOES) and the Advanced Very High Resolution Radiometer (AVHRR). The other method will use software modules from the Sensor Model Library to simulate the output of a number of different sensors from up-coming missions. These modules will be utilized during the second and third years when accurate data is required for testing the mature on-board processing system. The development of these modules will require close cooperation with the respective instrument teams.

Embedded On Board Processing System. The heart of the testbed is the Embedded On-Board Processing System (EOBPS). This component will be implemented from a combination of technologies. Standard Pentium class workstations will be utilized to execute the EVE system with minimal constraints. Additional hardware such as older generation processors, various embedded processor reference kits with either or both of the QNX or VxWorks real-time operating systems, and special purpose components will be used to simulate the constrained environment of on-board processing. While it will not be necessary to actually apply environmental hardening, the processing components must match those of real embedded space-based and UAV platforms. The ITSC will utilize other UAH and GHCC laboratory facilities to fabricate any special purpose hardware necessary to support the testbed. The EOBPS will execute on separate hardware from the Testbed Control System and Sensor Simulations to ensure that errors from those systems are isolated from the EOBPS.

Testbed Control. Additional utilities will be required to prepare the on-board processing testbed for experiments. Prior to testing any embedded modules the testbed must be calibrated to make sure that all external components are operating correctly. The sensor data must arrive at the on-board processing interface in the correct sequence and with proper timing. Experiment support files such as processing plans must be developed and tested external to the embedded system. Instrumentation of the testbed must be adapted for each experiment and checked independently. These and other related support utilities are essential to ensure that the testbed does not induce errors.

Comparative Technology Assessment

The availability of custom processing and data mining technologies on-board remote sensing platforms will enable real-time creation of specialized data products for direct transmission to users, as foreseen in NASA's Earth Science Vision. This new capability will support additional applications such as better natural hazard alerts and intelligent sensor and platform control.

The ITSC has been at the forefront of custom processing, feature extraction, and data mining technology for several years. Past research includes the development, component test, and system test of a ground-based data mining environment, in which selected data mining algorithms have been applied to Earth Science data and problems in a number of case studies. ADaM supports a number of research and production processing and data mining efforts in various configurations. The ADaM architecture, while using many advanced technologies, was not designed for use in an embedded on-board system environment. Using lessons learned in development of the existing ADaM system, the ITSC will provide an advanced information system capable of moving these types of applications, and others, into the real-time environment of space, at reduced risk, cost, and development time. ITSC's collaborations with Earth scientists in the development of on-board processing scenarios will serve to focus the proposed research on goals meaningful to both Information Technology and Earth Science applications.

Technology Readiness

As part of its software engineering practices, the ITSC routinely evaluates its research projects based upon NASA's Technology Readiness Levels. These range from 1 to 13 and represent a progression from initial research studies through flight hardware and software.

The current technology readiness level of ITSC custom processing and data mining technology for on-board processing is TRL 2 (Technology concept and/or application formulated.) While the ADaM system is relatively mature for its current uses in the ground-based environment, this system serves as only a preliminary analysis of this type of application for use in space. The EVE research will move this custom processing and data mining technology forward to TRL 5 over the course of three years. TRL 5 requires verification in a relevant environment. A laboratory testbed will be developed along with the processing prototype, in order to validate the new software architecture in a simulated on-board environment. The following sections describe the research and development required to reach each of these levels.

The first phase of this research, expected to take 1.5 years, will provide the proof-of-concept for feature extraction and data mining in an on-board environment. During this phase, technical feasibility of the project will be demonstrated with representative data in a laboratory testbed, with a primary focus on the implementation of key aspects of the EVE architecture. This implementation will include a small library of processing components, such as image processing and signal processing. A static assembly of processing components will be demonstrated on a representative hardware architecture. The results of this processing will be checked for correctness against the expected results. ADaM may be used in the verification effort.

The research phase required to develop a testable system is expected to take one year, with completion expected 2.5 years into the project. Demonstrations of the EVE system will follow the initial on-board processing scenario developed in collaboration with Earth Science researchers. This demonstration will use a processing plan, specific to the scenario, to assemble the required EVE modules. This test scenario will use the Embedded On-Board Processing System of the testbed, and a stream of real L0 data or simulated data from a planned mission from the testbed's Sensor Model Library.

The final six months of research will be focused on moving the EVE prototype toward verification in a real and relevant environment. At this point the full testbed and supporting systems will have been implemented, to support thorough testing of the prototype in an environment representative of an on-board processing system. The mechanism for introducing components, configuration, and data will be developed and tested in the target environment.

All of these activities will be undertaken with the goal of eventually integrating the processing system into an ongoing technology program such as the New Millennium Program. If, during the latter stages of this research project, a UAV flight is available, ITSC will reorient tasks as necessary to take advantage of this opportunity.

References

[1] Beaumont B., Helen Conover, and Sara Graves, 1996. "Information Systems Research at the Global Hydrology and Climate Center", American Institute of Aeronautics and Astronautics Spaces Programs and Technologies Conference, September 24 - 26, 1996.

[2] Conover H. and Sara J. Graves, 1999. "Promoting Science Data through Innovative Information Systems", American Geophysical Union, January 1999.

[3] Ramachandran R., H. Conover, S. Graves, K. Keiser, C. Pearson and J. Rushing, 1999. "A Next Generation Information System for Earth Science Data", The International Symposium on Optical Science, Engineering and Instrumentation, Denver, 1999.

[4] Graves S., 1998. "Automating the Process of Information Extraction in Digital Libraries", Panel Chair. IEEE Forum on Research and Technology Advances in Digital Libraries, April 22-24, 1998, Santa Barbara, CA.

[5] Keiser, K., J. Rushing, H. Conover and S. Graves, 1999. "Data Mining System Toolkit for Earth Science Data", Earth Observation and Geo-Spatial Web and Internet Workshop (EOGEO)-1999, Washington, Feb 9-11.

[6] Graves S., 1998. "Creating an Infrastructure for Data Warehousing and Mining", RCI-NASA Applications of Data Warehousing and Mining. Panel: Strategic Directions of Data Warehousing and Mining, April 20, 1998, Santa Barbara, CA.

[7] Ramachandran R. , Helen Conover, Sara Graves and Ken Keiser, 2000. "Algorithm Development and Mining (ADaM) System for Earth Science Applications", Second Conference on Artificial Intelligence, 80th AMS Annual Meeting, Long Beach, January, 2000.

[8] Hinke, Thomas H., J. Rushing, H. Ranganath and S. Graves, 1999. "Techniques and Experience in Mining Remotely Sensed Satellite Data", Artificial Intelligence Review: Issues on the Application of Data Mining, 2000.

[9] Hinke, T., J. Rushing, S. Kansal, S. Graves, H. Ranganath and E. Criswell, 1997. "Eureka Phenomena Discovery and Phenomena Mining System", 13th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography and Hydrology, February 1997.

[10] Nair, U.S., J. Rushing, R. Ramachandran, K. S. Kuo, S. Graves and R. M. Welch, 1999. "Detection of Cumulus Cloud Fields in Satellite Imagery", The International Symposium on Optical Science, Engineering, and Instrumentation, 18-23 July 1999, Denver, Colorado.

[11] Graves S., T. Hinke, H. Ranganath and J. Rushing, 1999. "Using Association Rules for Texture Characterization", Submitted to IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 23-25, 1999.

[12] Hinke, T., J. Rushing, H. Ranganath and S. Graves, 1997. "Target-Independent Mining for Scientific Data: Capturing Transients and Trends for Phenomena Mining", Proc. Third Int. Conf. On Data Mining (KDD-97), Newport Beach, CA, Aug. 14-17, 1997.

[13] Smith M., Bruce Beaumont, Susan McCoy, and Sara J. Graves, 1997. "Subsetting Data for EOSDIS", American Meteorological Society 13th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, February 2-7, 1997.

[14] Ramachandran R., M. Alshayeb, B. Beaumont, H. Conover, S. Graves, N. Hanish, X. Li, S. Movva, A. McDowell, and M. Smith, "Earth Science Markup Language", 17th Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, January 2001.